

## SEARCH REQUEST FORM

## Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched.

Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Point of Contact:

Jan Dalsal

Librarian-Physical Sciences

CM1 1E01 Tel: 308-4498

\*\*\*\*\*  
**STAFF USE ONLY**

Searcher: Jan  
 Searcher Phone #: 4458  
 Searcher Location: 118  
 Date Searcher Picked Up: 118  
 Date Completed: 119  
 Searcher Prep & Review Time: 10  
 Clerical Prep Time: 10  
 Online Time: 10

Type of Search	Vendors and cost where applicable
NA Sequence (#) <input checked="" type="checkbox"/>	STN _____
AA Sequence (#) <input checked="" type="checkbox"/>	Dialog _____
Structure (#) _____	Questel/Orbit _____
Bibliographic _____	Dr. Link _____
Litigation _____	Lexis/Nexis _____
Fulltext _____	Sequence Systems <input checked="" type="checkbox"/>
Patent Family _____	WWW/Internet _____
Other _____	Other (specify) _____



BASE COUNT 90 a 85 c 80 g 93 t  
ORIGIN

Query Match 100.0%; Score 348; DB 6; Length 348;  
Best Local Similarity 100.0%; Pred. No. 4 3e-103; Mismatches 0; Indels 0; Gaps 0;  
Matches 348; Conservative 0; MisMatches 0; Indels 0; Gaps 0;

Qy 1 atgccaacagacacgaggagccctgtggcccaagtcacaaatcggtttgtgtgg 60  
Jb 1 ATGCCAACAGATCACGAGCAGCCTGTTGCTCAATGG 60

Qy 61 gggtttttatgtgataccatattccggccattttgggtggatggatgg 120  
Jb 61 GGCTTGTGTACTATGCCACGCCATTGTAGGCGCTGAAACTAT 120

Qy 121 acaggagctgtgtggagatgtttttcccaaggccatcaactaaatgg 180  
Jb 121 ATGCCAACAGATCACGAGCAGCCTGTTGCTCAATGG 180

Qy 181 ctattttaaagttttggatggcttgcgtttttcccaaggccatcaactaaatgg 240  
Jb 181 CTGTTGAGCTTGTGGCATGTCAGTCAGTCAGTCAGTATGATAAC 240

Qy 241 ttcccttgccaggatccatgtggccacacatccaaactaaatgg 300  
Jb 241 TCTCTTGAGGAAAGGCCATCTAGAGGCACTTCAGTCAGTCAGTATGATAAC 300

Qy 301 gttagatcgaggatccatgtggccacacatccaaactaaatgg 348  
Jb 301 GTAGAGCAACATACAGAACCCGTCACGCACAGAACACTGA 348

RESULT 2  
AX08170 AX08170 348 bp DNA PAT 17-MAR-2001  
DEFINITION Sequence 1 from Patent WO011415.  
^CESSION AX08170  
VERSION AX08170.1 GI:13397083  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 348)  
AUTHORS Kavanaugh, W.M., Cen, H. and Lee, P.  
TITLE JOURNAL  
PATENT US  
PATENT NUMBER WO 01/1415-A 01-MAR-2001;  
CHIRON CORPORATION  
FEATURES Location/Qualifiers  
SOURCE /organism="Mus musculus"  
/db\_xref="taxon:10090"  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
1. .730  
/gene="Nr4"  
64. .411  
/gene="Nr4"  
/note="Erbb-4 ligand"  
/codon\_start=1  
/product="neuregulin-4 short isoform"  
/protein\_id="P421874.1"  
/db\_xref="GI:4550130"  
/translate="MTDHEQPGPRHRSFLNGGCVVIPNPSPFCRLENVYQAR  
CEEVFPPSSIPSESPNSAFLVAVLVLTLIAFLCFLCRGHQKASSVQEISLVE  
TNTRPRHSRHEH"  
ORIGIN  
BASE COUNT 82 a 102 c 87 g 77 t  
ORIGIN

Query Match 72.4%; Score 252; DB 6; Length 348;  
Best Local Similarity 82.8%; Pred. No. 1 2e-71; Mismatches 0; Indels 0; Gaps 0;  
Matches 288; Conservative 0; MisMatches 0; Indels 0; Gaps 0;

Qy 1 atgccaacagacacgaggagccctgtggcccaagtcacaaatcggtttgtgtgg 60  
Jb 1 ATGCCAACAGATCACGAGCAGCCTGTTGCTCAATGG 60

Qy 61 gggtttttatgtgataccatattccggccattttgggtggatggatgg 120  
Jb 61 GGCTTGTGTACTATGCCACGCCATTGTAGGCGCTGAAACTAT 120

Qy 121 acaggagctgtgtggagatgtttttcccaaggccatcaactaaatgg 180  
Jb 121 ATGCCAACAGATCACGAGCAGCCTGTTGCTCAATGG 180

Qy 1781 ctgtttaaatgtggacttggccatgtggccatgtggatggatgg 240  
Jb 181 CTGTTGAGCTTGTGGCATGTCAGTCAGTCAGTCAGTATGATAAC 240

Qy 241 ttcccttgccaggatccatgtggccacacatccaaactaaatgg 300  
Jb 241 TCTCTTGAGGAAAGGCCATCTAGAGGCACTTCAGTCAGTCAGTATGATAAC 300

Qy 301 gttagatcgaggatccatgtggccacacatccaaactaaatgg 348  
Jb 301 GTAGAGCAACATACAGAACCCGTCACGCACAGAACACTGA 348

RESULT 3  
AF083067 AF083067 750 bp mRNA ROD 03-JUN-1999  
DEFINITION Mus musculus neuregulin-4 short isoform (Nr4) mRNA, complete cds.  
ACCESSION AF083067  
VERSION AF083067.2 GI:4539580  
KEYWORDS house mouse.  
ORGANISM Mus musculus.  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 750)  
AUTHORS Harari, D., Tzahar, E., Romano, J., Shelly, M., Pierce, J.H., Andrews, G.C. and Yarden, Y.  
TITLE Neuregulin-4: a novel growth factor that acts through the Erbb-4 receptor tyrosine kinase.  
JOURNAL Oncogene 18 (17), 2681-2689 (1999)  
MEDLINE 99276098  
PUBLMED 10348342  
REFERENCE 2 (bases 1 to 750)  
AUTHORS Harari, D. and Yarden, Y.  
TITLE Direct Submission  
JOURNAL Submitted (09-Aug-1998) Biological Regulation, The Weizmann Institute of Science, Rehovot 76100, Israel  
COMMENT On Mar 30, 1999 this sequence version replaced g1:4530129.  
FEATURES Source  
SOURCE 1. .750  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
1. .730  
/gene="Nr4"  
64. .411  
/gene="Nr4"  
/note="Erbb-4 ligand"  
/codon\_start=1  
/product="neuregulin-4 short isoform"  
/protein\_id="P421874.1"  
/db\_xref="GI:4550130"  
/translate="MTDHEQPGPRHRSFLNGGCVVIPNPSPFCRLENVYQAR  
CEEVFPPSSIPSESPNSAFLVAVLVLTLIAFLCFLCRGHQKASSVQEISLVE  
TNTRPRHSRHEH"  
ORIGIN  
BASE COUNT 209 a 191 c 164 g 186 t  
ORIGIN

Query Match 72.4%; Score 252; DB 10; Length 750;  
Best Local Similarity 82.8%; Pred. No. 1 4e-71; Mismatches 0; Indels 0; Gaps 0;  
Matches 288; Conservative 0; MisMatches 0; Indels 0; Gaps 0;

Qy 1 atgccaacagacacgaggagccctgtggcccaagtcacaaatcggtttgtgtgg 60  
Jb 1 ATGCCAACAGATCACGAGCAGCCTGTTGCTCAATGG 60

Qy 61 gggtttttatgtgataccatattccggccattttgggtggatggatgg 120  
Jb 61 GGCTTGTGTACTATGCCACGCCATTGTAGGCGCTGAAACTAT 120

Qy 121 acaggagctgtgtggagatgtttttcccaaggccatcaactaaatgg 180  
Jb 121 ATGCCAACAGATCACGAGCAGCCTGTTGCTCAATGG 180

Qy 181 ctgtttaaatgtggacttggccatgtggccatgtggatggatgg 240  
Jb 181 CTGTTGAGCTTGTGGCATGTCAGTCAGTCAGTCAGTATGATAAC 240

Qy 241 ttcccttgccaggatccatgtggccacacatccaaactaaatgg 300  
Jb 241 TCTCTTGAGGAAAGGCCATCTAGAGGCACTTCAGTCAGTCAGTATGATAAC 300

Qy 301 gttagatcgaggatccatgtggccacacatccaaactaaatgg 348  
Jb 301 GTAGAGCAACATACAGAACCCGTCACGCACAGAACACTGA 348

ved Jan 9 06:07:28 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comp

Copyright (c) 1993 - 2000 compugen Ltd.

Bovine glial cell

Bovine glial cell  
Human neuregulin  
Glial growth factor  
Bovine glial growth

protein - protein search, using sw model  
on: January 8, 2002, 10:49:43 : Search time 23.58 seconds

361.257 Million cell updates/sec

fluence: 1 MPTTHEEPCGGPSHSKFCLNG.....YDINVETSTSSTAHHSHEQH 115  
ring table: BLOSUM62 ^ GCGCTAT^G^E

searched: 522463 seqs, 74073290 residues  
cal number of hits satisfying chosen parameters: 522463

minimum DB seq length: 0  
maximum DB seq length: 200000000

Maximum Match 100%  
Listing first 45 summaries  
\* CONCERN 1-01 \*

```

1: /$SUS8/gcdata/targeneseq/geneseq/PA1980.DAT: *
2: /$STD8/gcdata/targeneseq/geneseq/PA1980.DAT: *
3: /$STD8/gcdata/targeneseq/geneseq/PA1982.DAT: *
4: /$STD8/gcdata/targeneseq/geneseq/PA1983.DAT: *

```

```

6: //SRS8 gedata geneseq geneseqp AA1985.DAT *
7: //SRS8 gedata geneseq geneseqp AA1985.DAT *
8: //SRS8 gedata geneseq geneseqp AA1987.DAT *

```

```

10: /SIDS8/gcadata/geneseq/geneseq/geneseqP/A11980.DAT;
11: /SIDS8/gcadata/geneseq/geneseqP/A11980.DAT;
12: /SIDS8/gcadata/geneseq/geneseqP/A11981.DAT;
13: /SIDS8/gcadata/geneseq/geneseqP/A11981.DAT;

```

```

14: /SDSB8/eggdata/geneseq/geneseq/An1994.DAT
15: /SDSB8/eggdata/geneseq/geneseq/An1994.DAT
16: /SDSB8/eggdata/geneseq/geneseq/An1995.DAT
17: /SDSB8/eggdata/geneseq/geneseq/An1996.DAT

```

```

19: /SIDS8/godata/geneseed/geneseq/AA1998.DAT
20: /SIDS8/godata/geneseed/geneseq/AN2000.DAT
21: /SIDS8/godata/geneseed/geneseq/AN2001.DAT

```

Pred. No. is the number of results predicted by chance to score greater than or equal to the score of the result being derived by analysis of the total score distribution.

## SUMMARIES

### Query

### § sult

1	625	100.0	115	22	AAX97697	Human
2	489	78.2	115	22	AAX97698	Mouse
3	307	20.2	115	22	AAX97699	Monkey

	4	5	6	7
140	22.4	602	15	ANR51705
140	22.4	18	ANR52900	Neurin
140	22.4	1070	22	ANR48099
136	21.8	696	22	ANR97619

9	136	21.8	720	20
10	132.5	21.2	ANR97618	Human
11	132.5	21.2	ANR46916	GCF
			ANR55657	GCF222

## SUMMARIES

DIGIMENE

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

XX EGFR2; epidermal growth factor H2; EGFR2 protein modulated disorder; KW cancer; peripheral neuropathy; amyotrophic lateral sclerosis; therapy; KW Alzheimer's disease; Parkinson's disease; Huntington's disease; KW ischaemic stroke; brain injury; acute spinal cord injury; infection; KW nervous system injury; multiple sclerosis; dementia; epilepsy; KW peripheral nerve injury; acoustic trauma; human.



GenCore version: 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model  
on: January 8, 2002, 10:58:38 ; Search time 22.24 Seconds  
(without alignments)

75.354 Million cell updates/sec  
ximum DB seq length: 0  
st-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

tabase : SPTRIMBL17:  
rfect score: 625  
fluence: 1 MPTDHEEPCGPSHSKSFCLNG.....YDINLVETSSSAHHSHEQH 115  
oring table: BLOSUM62  
GapOp 10.0 , Gapext 0.5  
arched: 473505 seqs, 146272329 residues  
,1 number of hits satisfying chosen parameters: 473505  
nimum DB seq length: 0  
ximum DB seq length: 200000000  
st-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

1: SPTRIMBL17:  
2: SP\_archea:  
3: SP\_bacteria:  
4: SP\_fungi:  
5: SP\_invertebrate:  
6: SP\_mammal:  
7: SP\_mhc:  
8: SP\_organelle:  
9: SP\_phage:  
10: SP\_plant:  
11: SP\_rhodent:  
12: SP\_virus:  
13: SP\_vertebrate:  
14: SP\_unclassified:  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

sult	Score	Query	Length	DB	ID	Description
1	130.5	20.9	461	11	035947	035947 metocricetu
2	128.5	20.6	1214	6	098D00	09bdq0 sus scrofa
3	123.5	19.8	298	11	09ESA9	09esa9 rattus norv
4	123.5	19.8	695	11	09ESB0	09esb0 rattus norv
5	122.5	19.6	162	11	061521	061521 mus musculu
6	118.5	19.0	162	11	09Z0L5	092015 rattus norv
7	117.5	18.8	127	12	09Q9F3	09q9f3 ectromelia
8	116	18.6	700	11	09ESB1	09esb1 rattus norv
9	116	18.6	782	11	09ESA5	09esa5 rattus norv
10	115	18.4	140	12	089756	089756 variola vir
11	113.5	18.2	127	12	041506	041506 been 58058
12	113.5	18.2	140	12	057166	057166 vaccinia vi
13	113.5	18.2	140	12	09JFH4	09Jfh4 vaccinia vi
14	112	17.9	140	12	089066	089066 variola vir
15	111.5	17.8	140	12	06607	08607 vaccinia vi
16	106.5	17.5	138	12	P87605	P87605 cowpox viru
17	109.5	17.5	241	6	Q07112	Q07112 bos taurus
18	108.5	17.4	125	12	Q41504	Q41504 cowpox viru
19	107.5	17.2	111	12	Q9ESAB	Q9esab rattus norv

RESULT	1
ID	035947
AC	035947;
DT	01-JAN-1998 (TREMBLrel. 05, Created)
DT	01-JUN-2001 (TREMBLrel. 05, Last sequence update)
DE	PRO-NEUREGULIN-1, ISOFORM ALPHA 2B PRECURSOR.
OS	Mesocricetus auratus (Golden hamster).
OC	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.
OX	NCBI_TAXID=10036;
RN	[1] SEQUENCE FROM N.A. (ISOFORM ALPHA2B), AND SEQUENCE OF 64-81.
RC	TISSUE=EMBRYO;
RX	MEDLINE=98196996; PUBMED=9537646;
RA	Velasco J.A., Reijo E., Avila M.A., Notario V.;
RT	"Secretion of neu differentiation factor-like polypeptides by capping transformed fibroblasts: cloning and characterization of Syrian hamster neurogulin cDNAs."
RT	Mol. Carcinog. 21:156-163 (1998).
RL	FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
CC	- RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS, RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND ACTIVATION OF THE ERBB RECEPTEORS. MAY PLAY AN IMPORTANT ROLE IN PROVIDING GROWTH ADVANTAGE IN NEOPLASTIC CELLS.
CC	- SUBUNIT: THE CYTOSLIC DOMAIN INTERACTS WITH THE LIM DOMAIN REGION OF LIMK1 (BY SIMILARITY).
CC	- SUBCELLULAR LOCATION: EXISTS AS TYPE I MEMBRANE PROTEIN AND AS A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC	- ALTERNATIVE PRODUCTS: DIFFERENT ISOFORMS ARE PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ISOFORM ALPHA2B/CLONE PM3.
CC	- TISSUE SPECIFICITY: EXPRESSED AT HIGHER LEVEL AFTER NEOPLASTIC TRANSFORMATION OF CELLS.
CC	- DOMAIN: THE CYTOSLIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN DIMERIZATION (BY SIMILARITY).

Kavanaugh WM, Cen H, Lee P;  
WPI; 2001-211304/21.  
P-PSDB; AAY91297.

Novel epidermal growth factor H2 (EGFH2) polypeptides and  
polynucleotides encoding them useful for treating EGF2  
protein-modulated disorders such as cancers, Alzheimer's disease,  
ischaemic stroke and epilepsy -

KW EGFH2; epidermal growth factor H2; EGFH2 protein modulated disorder;  
 KW cancer; peripheral neuropathy; amyotrophic lateral sclerosis; therapy;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KW ischaemic stroke; brain injury; acute spinal cord injury; infection;  
 KW nervous system injury; multiple sclerosis; dementia; epilepsy;  
 KW peripheral nerve injury; acoustic trauma; mouse; ds.  
 XX MUS sp.  
 XX OS  
 XX  
 FH Location/Qualifiers  
 FT 1..348  
 FT /\*tag\_a  
 FT /product\_ "EGFH2"  
 PN WO200114415-A2.  
 XX  
 PD 01-MAR-2001.  
 XX  
 PR 15-AUG-2000; 2000WO-US22326.  
 XX 20-AUG-1999; 99US-0149986.  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PT Kavanaugh WM, Cen H, Lee P;  
 XX  
 DR WPI; 2001-211304/21.  
 DR P-PSDB; AAY97698.  
 XX  
 PT Novel epidermal growth factor H2 (EGFH2) polypeptides and  
 PT polynucleotides encoding them useful for treating EGFH2  
 PT protein-modulated disorders such as cancers, Alzheimer's disease,  
 PT ischaemic stroke and epilepsy  
 XX  
 PS Claim 2; Fig 3; 48pp; English.  
 XX  
 CC This sequence encodes the mouse epidermal growth factor H2 (EGFH2)  
 CC protein of the invention. The EGFH2 DNA and protein sequences are useful  
 CC for modulating the amount of EGFH2 protein in a subject, or treating an  
 CC EGFH2 protein modulated disorder. The DNA or its antisense construct,  
 CC ribozyme or a retroviral vector comprising a promoter are also useful  
 CC for treating breast, prostate, pancreatic, oral or ovarian cancers,  
 CC peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease, ischaemic stroke,  
 CC brain injury, acute spinal cord injury, nervous system injury, multiple  
 CC sclerosis, infection, dementia, epilepsy, peripheral nerve injury,  
 CC acoustic trauma and tissue wound. The DNA and protein are also useful for  
 CC providing trophic support for cells in a patient by implanting cells  
 CC expressing them encapsulated in a semipermeable membrane. EGFH2  
 CC polypeptides can be used to screen peptide libraries. Antibodies which  
 CC bind to EGFH2 and/or variant polypeptides can be used in diagnosing and  
 CC determining the prognosis of cancer, tumour progression,  
 CC hyperproliferative cell growth or accompanying biological, and physical  
 CC manifestations.  
 XX  
 Sequence 348 BP; 82 A; 102 C; 87 G; 77 T; 0 other;  
 SQ  
 Query Match 72.4%; Score 252; DB 22; Length 348;  
 Best Local Similarity 82.8%; Pred. No. 2.8e-74;  
 Matches 288; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
 Matches 288;  
 1 atgcccacacatcaccggaggccctgtggcccaatgcgtttgcgtatgg 60  
 1 atgcacaaatggatccacggaggccctgtggcccaatgcgtatgg 60  
 61 gggtttttatgtatccatctatccacggccatgtttgtggatggaaactat 120  
 61 ||||||||||||||||||||||||||||||||||||||||||||||||||||| 120  
 61 gggtttttatgtatccatctatccacggccatgtttgtggatggaaactat 120  
 121 acggaggatcggtgtggagggttttcccggttcacgtccaaactaaaggtaac 180  
 121 acggaggatcggtgtggagggttttcccggttcacgtccaaactaaaggtaac 180

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model  
on: January 8, 2002, 10:58:58 ; search time 11.71 seconds  
maximum DB seq length: 0  
st-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

file: US-09-640-041-4  
reflect score: 625  
sequence: 1 MPIDHEPCGFSHKSFCLNG.....YDINLVETSSTSAAHSHEQH 115

oring table: BLOSUM62  
Gapop 10.0 , gapext 0.5  
arched: 100059 seqs, 36664827 residues  
tal number of hits satisfying chosen parameters: 100059  
nimum DB seq length: 0  
ximum DB seq length: 200000000

MAXIMUM Match 0%  
Listing first 45 summaries

### SUMMARIES

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1  
NRG4\_MOUSE STANDARD; PRT; 115 AA.  
ID NRG4\_MOUSE  
AC Q9WTA4;  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PRO-NEUREGULIN-4, SHORT ISOFORM (PRO-NRG4) (CONTAINS: NEUREGULIN-4 (NRG-4)).  
GN NRG4.  
OS Mus musculus (Mouse).  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=99276098; PubMed=10348342;  
RA Harari D., Tahir E., Romano J., Shelly M., Pierce J.H., Andrews G.C., RA Yarden Y.;  
RT "Neuregulin-4: a novel growth factor that acts through the Erbb-4 receptor tyrosine kinase";  
RL Oncogene 18:2281-2287(1999).  
CC -1- FUNCTION: LOW AFFINITY LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS, RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND ACTIVATION OF THE ERBB RECEPTORS. DOES NOT BIND TO THE ERBB1, ERBB2 AND ERBB3 RECEPTORS.  
CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS MAY BE PRODUCED BY ALTERNATIVE SPlicing.  
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PANCREAS; WEEKLY EXPRESSED IN MUSCLE.  
CC -1- DOMAIN: THE CYTOPLASMATIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN DIMERIZATION (BY SIMILARITY).  
CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE DOMAIN (BY SIMILARITY).  
CC -1- DOMAIN: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR (BY SIMILARITY).  
CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.

006175 rattus norv P35070 homo sapien  
P35070 homo sapien  
09tcc5 bos taurus  
P45442 caenorhabdi P10040 drosophila  
P16819 homo sapien P48030 mus musculus  
P01134 rattus norvegicus  
P05926 mus musculus P21955 mus musculus  
P28858 macaca mulatta  
P28858 macaca mulatta

### ALIGNMENTS

title: US-09-640-041-4  
reflect score: 625  
sequence: 1 MPIDHEPCGFSHKSFCLNG.....YDINLVETSSTSAAHSHEQH 115

oring table: BLOSUM62  
Gapop 10.0 , gapext 0.5  
arched: 100059 seqs, 36664827 residues  
tal number of hits satisfying chosen parameters: 100059  
nimum DB seq length: 0  
ximum DB seq length: 200000000

MAXIMUM Match 0%  
Listing first 45 summaries

### SUMMARIES

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1  
NRG4\_MOUSE STANDARD; PRT; 115 AA.  
ID NRG4\_MOUSE  
AC Q9WTA4;  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PRO-NEUREGULIN-4, SHORT ISOFORM (PRO-NRG4) (CONTAINS: NEUREGULIN-4 (NRG-4)).  
GN NRG4.  
OS Mus musculus (Mouse).  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=99276098; PubMed=10348342;  
RA Harari D., Tahir E., Romano J., Shelly M., Pierce J.H., Andrews G.C., RA Yarden Y.;  
RT "Neuregulin-4: a novel growth factor that acts through the Erbb-4 receptor tyrosine kinase";  
RL Oncogene 18:2281-2287(1999).  
CC -1- FUNCTION: LOW AFFINITY LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS, RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND ACTIVATION OF THE ERBB RECEPTORS. DOES NOT BIND TO THE ERBB1, ERBB2 AND ERBB3 RECEPTORS.  
CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS MAY BE PRODUCED BY ALTERNATIVE SPlicing.  
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PANCREAS; WEEKLY EXPRESSED IN MUSCLE.  
CC -1- DOMAIN: THE CYTOPLASMATIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN DIMERIZATION (BY SIMILARITY).  
CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE DOMAIN (BY SIMILARITY).  
CC -1- DOMAIN: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR (BY SIMILARITY).  
CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

title: US-09-640-041-4  
reflect score: 625  
sequence: 1 MPIDHEPCGFSHKSFCLNG.....YDINLVETSSTSAAHSHEQH 115

oring table: BLOSUM62  
Gapop 10.0 , gapext 0.5  
arched: 100059 seqs, 36664827 residues  
tal number of hits satisfying chosen parameters: 100059  
nimum DB seq length: 0  
ximum DB seq length: 200000000

MAXIMUM Match 0%  
Listing first 45 summaries

### SUMMARIES

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1  
NRG4\_MOUSE STANDARD; PRT; 115 AA.  
ID NRG4\_MOUSE  
AC Q9WTA4;  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PRO-NEUREGULIN-4, SHORT ISOFORM (PRO-NRG4) (CONTAINS: NEUREGULIN-4 (NRG-4)).  
GN NRG4.  
OS Mus musculus (Mouse).  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=99276098; PubMed=10348342;  
RA Harari D., Tahir E., Romano J., Shelly M., Pierce J.H., Andrews G.C., RA Yarden Y.;  
RT "Neuregulin-4: a novel growth factor that acts through the Erbb-4 receptor tyrosine kinase";  
RL Oncogene 18:2281-2287(1999).  
CC -1- FUNCTION: LOW AFFINITY LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS, RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND ACTIVATION OF THE ERBB RECEPTORS. DOES NOT BIND TO THE ERBB1, ERBB2 AND ERBB3 RECEPTORS.  
CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS MAY BE PRODUCED BY ALTERNATIVE SPlicing.  
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PANCREAS; WEEKLY EXPRESSED IN MUSCLE.  
CC -1- DOMAIN: THE CYTOPLASMATIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN DIMERIZATION (BY SIMILARITY).  
CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE DOMAIN (BY SIMILARITY).  
CC -1- DOMAIN: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR (BY SIMILARITY).  
CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WinPatin (Genentech)

CURRENT APPLICATION DATA:  
 CURRENT APPLICATION DATE:  
 REFERENCE NUMBER: US/08/899, 437  
 FILING DATE: 24-JUL-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conley, Deirdre L.  
 REGISTRATION NUMBER: 36,487  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-2066  
 TELEFAX: 650/952-9881

SEQUENCE CHARACTERISTICS:  
 LENGTH: 2538 base pairs  
 TYPE: Nucleic Acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear

SEQUENCE FOR SEQ ID NO: 1:  
 NAME/KEY: mouse NRG3 nucleic acid  
 LOCATION: 1-2538  
 IDENTIFICATION METHOD:  
 OTHER INFORMATION:  
 899-437-1

SEQUENCE CHARACTERISTICS:  
 LENGTH: 2538 base pairs  
 TYPE: Nucleic Acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear

FEATURE:  
 NAME/KEY: mouse NRG3 nucleic acid  
 LOCATION: 1-2538  
 IDENTIFICATION METHOD:  
 OTHER INFORMATION:  
 US-09-126-121-1

Query Match 8.6%; Score 29.8; DB 3; length 2538;  
 Best Local Similarity 55.0%; Pred. No. 1.3; Mismatches 0; Indels 6; Gaps 1;

4 ccaacagatcacgaagagccgtggccggcggatccggatgggggg 63  
 141 CGATCTGAGCACTCAACCCCTGAGACAGGACTCTGGTCAATGAGGT 1200

Qy 4 ccaacagatcacgaagagccgtggccggcggatccggatgggggg 63  
 Db 1141 CGATCTGAGCACTCAACCCCTGAGACAGGACTCTGGTCAATGAGGT 1200

Qy 64 ctttgttatgtataactactatccggccattt----ttaggtgggtggaaac 117  
 Db 1141 CGATCTGAGCACTCAACCCCTGAGACAGGACTCTGGTCAATGAGGT 1200

Qy 118 tatacaggactctgtggaaagggttttc 148  
 Db 1261 TACCAAGGAGTCGGTGTGATCAATTCTGC 1291

RESULT 4  
 US-08-482-918-47  
 ; Sequence 47, Application US/08482918  
 ; Patent No. 6207417

GENERAL INFORMATION:  
 APPLICANT: Zeebo, Krisztina M.  
 APPLICANT: Boselman, Robert A.  
 APPLICANT: Singgs, Sidney V.  
 APPLICANT: Martin, Francis H.  
 TITLE OF INVENTION: Stem Cell Factor  
 NUMBER OF SEQUENCES: 104

CORRESPONDENCE ADDRESS:  
 ADDRESSE: Marshall, O'Toole, Gorstein, Murray & Borun  
 STREET: 6300 Sears Tower, 233 South Wacker Drive  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: United States of America  
 ZIP: 60606-6402

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/482, 918  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
 NAME: Clough, David W.  
 REGISTRATION NUMBER: 36,107

SOFTWARE: WinPatin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09-126, 121  
 FILING DATE: 30-Jul-1998  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conley, Deirdre L.  
 REGISTRATION NUMBER: 36,487  
 REFERENCE/DOCKET NUMBER: P1084R1D1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-2066  
 TELEFAX: 650/952-9881

HS-09-640-041-4.FBF

;Note: sequence extracted from NCBI backbone (NCBIN:127787, NCBIPI:127788)

C;Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 22.4%; Score 140; DB 2; Length 602; Best Local Similarity 27.1%; Pred No. 8.3e-07; Matches 35; Conservative 21; Mismatches 51; Indels 22; Gaps 4;

Qy 3 TDHEEPCGGSHKSCLNGGLCYVPTIPSP--FRCVENVYGARCEEV----- 48 Db 135 TSHLTKCDIKRQKATCVCNGECYKWDLPRPVLRCRCNEFTGDRCONVYMASFYKHLGI 194 Qy 49 -FLPSSQIQTNSNIEFAFVALVNLVLIIGAFFLCL---RKGHFORASSVQYDINLV 101 Db 195 EFMRREELVKRVLITGICILLYVNGMCVWAY-CKTKKQRKKLHDLRQLRSERNV 253

Qy 102 ETSISAH 110 Db 254 MMANGPHH 262

RESULT 3 neu differentiation factor - human  
C;Species: Homo sapiens (man) #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C;Accession: I38404 R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14: 1909-1919, 1994  
A;Title: structural and functional aspects of the multiplicity of Neu differentiation fa  
A;Reference number: A56210; MUID:94158863  
A;Status: preliminary; translated from GB/EMBL/DBBJ  
A;Molecule type: mRNA  
A;Residues: 1462 <RES>  
A;Cross-references: GDB:132656; OMIM:142445  
A;Map position: 8p22-8p11  
C;Superfamily: unassigned EGF-related proteins; EGF homology  
C;Keywords: alternative splicing; glycoprotein F;182-221/Domain: EGF homology <EGF>

Query Match 20.2%; Score 126.5; DB 2; Length 462; Best Local Similarity 25.0%; Pred. No. 1.6e-05; Matches 32; Conservative 25; Mismatches 46; Indels 25; Gaps 6; C;Species: Homo sapiens (man) #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C;Accession: I38404 R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14: 1909-1919, 1994  
A;Title: structural and functional aspects of the multiplicity of Neu differentiation fa  
A;Reference number: A56210; MUID:94158863  
A;Status: preliminary; translated from GB/EMBL/DBBJ  
A;Residues: 1462 <RES>  
A;Cross-references: EMBL:002326; NID:9408402; PIDN:AAA19951.1; PID:9408403

Query Match 20.2%; Score 126.5; DB 2; Length 462; Best Local Similarity 25.0%; Pred. No. 1.6e-05; Matches 32; Conservative 25; Mismatches 46; Indels 25; Gaps 6;

Qy 3 TDHEEPCGGSHKSCLNGGLCYVPTIPSP--FRCVENVYGARCEEVFLPGSSIQKS 59 Db 176 TSHLVKCAEKEKTFCVNGECYKWDLNSPNSRLCKCQPGFTGARCTE-NVP-MKVQNE 233

Qy 60 NLFFAFVALVLT---LIGAFYFLC-----RKGHFORASSVQYDINLV 102

Db 234 KARELYQKRVLTGICILLYVNG---IMCVVAVCKTKKQRKKLHDLRQLRSERNMM 290  
Qy 103 TSSISAH 110 Db 291 NIANGPHH 298

RESULT 5 heresulin precursor, splice form beta-2 - human

C;Species: Homo sapiens (man) #sequence\_revision 31-Dec-1993 #text\_change 17-Nov-2000  
C;Accession: I38407 R;Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yee Science 256, 1205-1210, 1992  
A;Title: Identification of heresulin, a specific activator of p185(erbb2).  
A;Reference number: A43273; MUID:92271253  
A;Accession: C43273  
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A;Molecule type: mRNA  
A;Residues: 1-637 <HOL>  
R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A;Title: structural and functional aspects of the multiplicity of Neu differentiation fa  
A;Reference number: A56210; MUID:94158863  
A;Accession: I38407  
A;Status: preliminary; translated from GB/EMBL/DBBJ  
A;Molecule type: mRNA  
A;Residues: 119-406 <RES>  
A;Cross-references: EMBL:002329; NID:9408408; PIDN:AAA19954.1; PID:9408409  
C;Genetics:  
A;Gene: GDB:HGL  
A;Cross-references: GDB:132656; OMIM:142445  
A;Map position: 8p22-8p11  
C;Superfamily: unassigned EGF-related proteins; EGF homology  
C;Keywords: alternative splicing F;182-221/Domain: EGF homology <EGF>

J. Biol. Chem. 268, 18407-18410, 1993  
A;Title: Characterization of a breast cancer cell differentiation factor that specifies  
A;Reference number: A48498; MUID:93366731  
A;Accession: A48498  
A;Molecule type: protein  
A;Residues: 20-21, 'X', 23-24, 'XX', 27-28 <CUT>  
R;Peleš, E.; Bacus, S.S.; Koski, R.A.; Lu, H.S.; Wen, D.; Ogden, S.G.; Levy, R.B.; Y Cell 69, 205-216, 1992  
A;Title: Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that is  
A;Reference number: A38155; MUID:92208945  
A;Accession: A38155  
A;Molecule type: protein  
A;Residues: 'X', 15-16, 'X', 18-20, 'RG', 23-24, 'GP', 27, 'E', 29, 'XP', 32-36 <PEL>  
A;Note: sequence extracted from NCBI backbone (NCBIPI:91347)  
C;genetics:  
A;Gene: GDB:HGL  
A;Cross-references: GDB:132656; OMIM:142445  
A;Map position: 8p22-8p11  
C;Superfamily: unassigned EGF-related proteins; EGF homology  
C;Keywords: alternative splicing; glycoprotein F;182-221/Domain: EGF homology <EGF>

RESULT 4 heresulin precursor, splice form alpha - human  
C;Species: Homo sapiens (man) #sequence\_revision 31-Dec-1993 #text\_change 11-Jan-2000  
C;Accession: A43273, A48498; A38155 R;Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yee Science 256, 1205-1210, 1992  
A;Title: Identification of heresulin, a specific activator of p185(erbb2).  
A;Reference number: A43273; MUID:92271253  
A;Accession: C43273  
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A;Molecule type: mRNA  
A;Residues: 1-637 <HOL>  
R;Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A;Title: structural and functional aspects of the multiplicity of Neu differentiation fa  
A;Reference number: A56210; MUID:94158863  
A;Accession: I38407  
A;Status: preliminary; translated from GB/EMBL/DBBJ  
A;Molecule type: mRNA  
A;Residues: 119-406 <RES>  
A;Cross-references: EMBL:002329; NID:9408408; PIDN:AAA19954.1; PID:9408409  
C;Genetics:  
A;Gene: GDB:HGL  
A;Cross-references: GDB:132656; OMIM:142445  
A;Map position: 8p22-8p11  
C;Superfamily: unassigned EGF-related proteins; EGF homology  
C;Keywords: alternative splicing F;182-221/Domain: EGF homology <EGF>

Query Match 19.8%; Score 124; DB 2; Length 637; A;Note: sequence extracted from NCBI backbone (NCBIN:103250, NCBIPI:103250)  
R;Culouscou, J.M.; Plowman, G.D.; Carlton, G.W.; Green, J.M.; Shoyab, M.



C C  
 R EMBL; AF083067; AAD21874; -.  
 R InterPro; IPR00561; EGF-like.  
 R InterPro; IPR00336; EGF\_1..  
 R Pfam; PF00008; EGF; 1.  
 R PRINTS; PRO0009; EGFGE.  
 R SMART; SM00181; EGF; 1.  
 R PROSITE; PS00022; EGF\_1; 1.  
 R PROSITE; PS01185; EGF\_2; FALSE\_NEG.  
 R Growth factor; EGF-like domain; Glycoprotein; Transmembrane;  
 W Multigene family; Alternative splicing.  
 T CHAIN 1 115 PRO-NEUREGULIN-4, MEMBRANE-BOUND FORM.  
 T DOMAIN 1 61 NEUREGULIN-4.  
 T DOMAIN 1 62 EXTRACELLULAR (POTENTIAL).  
 T TRANSMEM 63 83 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 T DOMAIN 84 115 CYTOPLASMIC (POTENTIAL).  
 T DOMAIN 5 46 EGF-LIKE.  
 T DISULFID 9 23 BY SIMILARITY.  
 T DISULFID 17 34 BY SIMILARITY.  
 T DISULFID 36 45 BY SIMILARITY.  
 T CARBOHYD 39 39 N-LINKED (GLCNAC : .) (POTENTIAL).  
 T CARBOHYD 60 60 N-LINKED (GLCNAC : .) (POTENTIAL).  
 Q SEQUENCE 115 AA; 1743 MW; 989AEB376F857B49 CRC64;  
 DR EMBL; AF010130; AAB70914; 1; -.  
 DR MGD; MGI:1057165; Nrr3.  
 DR InterPro; IPR00561; EGF-like.  
 DR InterPro; IPR002154; Neuregulin.  
 DR Pfam; PF00008; EGF; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR Growth factor; EGF-like domain; Transmembrane; Multigene family  
 DR PRO-NEUREGULIN-3, MEMBRANE-BOUND FORM.  
 FT CHAIN 1 713 PRO-NEUREGULIN-3.  
 FT DOMAIN 1 361 NEUREGULIN-3.  
 FT DOMAIN 363 383 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 383 384 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 FT DOMAIN 105 287 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 288 331 SER/THR-RICH.  
 FT DOMAIN 13 21 EGF-LIKE.  
 FT DOMAIN 26 34 POLY-ALA.  
 FT DOMAIN 127 135 POLY-THR.  
 FT DOMAIN 250 253 POLY-ALA.  
 FT DOMAIN 254 263 POLY-SER.  
 FT DOMAIN 264 267 POLY-THR.  
 FT DISULFID 292 306 BY SIMILARITY.  
 FT DISULFID 300 319 BY SIMILARITY.  
 FT DISULFID 321 330 BY SIMILARITY.  
 FT DISULFID 321 330 BY SIMILARITY.  
 SQ SEQUENCE 713 AA; 77369 MW; 997D1D5E7FC8DCFO CRC64;  
 DR SEQUENCE FROM N.A.  
 DR TISSUE-BRAIN;  
 RX MEDLINE:97420720; PubMed=9275162;  
 RA Zhang D., Sliwowski M.X., Mark M., Frantz G., Akita R., Sun Y.,  
 RA Hillen K., Crowley J., Brush J., Godowski P.J.;  
 RT "Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that  
 binds and activates ErbbA." Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).  
 RL - FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR.  
 CC BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR,  
 CC ERBB2 OR ERBB3 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN SYMPATHETIC, MOTOR, AND SENSORY  
 CC NEURONS.  
 CC -1- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS E11. IN E13 EMBRYOS, DETECTED IN  
 CC THE BRAIN, SPINAL CORD, TRIGEMINAL, VESTIBULAR-COCHLEAR, AND  
 CC

CC SPINAL GANGLIA. IN ADULTS, EXPRESSED IN SPINAL CORD, AND NUMEROUS  
 CC BRAIN REGIONS.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 CC DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 CC DOMAIN (BY SIMILARITY).  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 CC FORM (BY SIMILARITY).  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is not modified  
 CC and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL; AF010130; AAB70914; 1; -.  
 DR MGD; MGI:1057165; Nrr3.  
 DR InterPro; IPR00561; EGF-like.  
 DR InterPro; IPR002154; Neuregulin.  
 DR Pfam; PF00008; EGF; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR Growth factor; EGF-like domain; Transmembrane; Multigene family  
 DR PRO-NEUREGULIN-3, MEMBRANE-BOUND FORM.  
 FT CHAIN 1 713 PRO-NEUREGULIN-3.  
 FT DOMAIN 1 361 NEUREGULIN-3.  
 FT DOMAIN 363 383 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 383 384 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 FT DOMAIN 105 287 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 288 331 SER/THR-RICH.  
 FT DOMAIN 13 21 EGF-LIKE.  
 FT DOMAIN 26 34 POLY-ALA.  
 FT DOMAIN 127 135 POLY-THR.  
 FT DOMAIN 250 253 POLY-ALA.  
 FT DOMAIN 254 263 POLY-SER.  
 FT DOMAIN 264 267 POLY-THR.  
 FT DISULFID 292 306 BY SIMILARITY.  
 FT DISULFID 300 319 BY SIMILARITY.  
 FT DISULFID 321 330 BY SIMILARITY.  
 SQ SEQUENCE 713 AA; 77369 MW; 997D1D5E7FC8DCFO CRC64;  
 DR SEQUENCE FROM N.A.  
 DR TISSUE-BRAIN;  
 RX MEDLINE:97420720; PubMed=9275162;  
 RA Zhang D., Sliwowski M.X., Mark M., Frantz G., Akita R., Sun Y.,  
 RA Hillen K., Crowley J., Brush J., Godowski P.J.;  
 RT "Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that  
 binds and activates ErbbA." Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).  
 RL - FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR.  
 CC BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR,  
 CC ERBB2 OR ERBB3 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN SYMPATHETIC, MOTOR, AND SENSORY  
 CC NEURONS.  
 CC -1- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS E11. IN E13 EMBRYOS, DETECTED IN  
 CC THE BRAIN, SPINAL CORD, TRIGEMINAL, VESTIBULAR-COCHLEAR, AND  
 CC

CC SPINAL GANGLIA. IN ADULTS, EXPRESSED IN SPINAL CORD, AND NUMEROUS  
 CC BRAIN REGIONS.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 CC DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 CC DOMAIN (BY SIMILARITY).  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 CC FORM (BY SIMILARITY).  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is not modified  
 CC and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL; AF010130; AAB70914; 1; -.  
 DR MGD; MGI:1057165; Nrr3.  
 DR InterPro; IPR00561; EGF-like.  
 DR InterPro; IPR002154; Neuregulin.  
 DR Pfam; PF00008; EGF; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR Growth factor; EGF-like domain; Transmembrane; Multigene family  
 DR PRO-NEUREGULIN-3, MEMBRANE-BOUND FORM.  
 FT CHAIN 1 713 PRO-NEUREGULIN-3.  
 FT DOMAIN 1 361 NEUREGULIN-3.  
 FT DOMAIN 363 383 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 383 384 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 FT DOMAIN 105 287 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 288 331 SER/THR-RICH.  
 FT DOMAIN 13 21 EGF-LIKE.  
 FT DOMAIN 26 34 POLY-ALA.  
 FT DOMAIN 127 135 POLY-THR.  
 FT DOMAIN 250 253 POLY-ALA.  
 FT DOMAIN 254 263 POLY-SER.  
 FT DOMAIN 264 267 POLY-THR.  
 FT DISULFID 292 306 BY SIMILARITY.  
 FT DISULFID 300 319 BY SIMILARITY.  
 FT DISULFID 321 330 BY SIMILARITY.  
 SQ SEQUENCE 713 AA; 77369 MW; 997D1D5E7FC8DCFO CRC64;  
 DR SEQUENCE FROM N.A.  
 DR TISSUE-BRAIN;  
 RX MEDLINE:97420720; PubMed=9275162;  
 RA Zhang D., Sliwowski M.X., Mark M., Frantz G., Akita R., Sun Y.,  
 RA Hillen K., Crowley J., Brush J., Godowski P.J.;  
 RT "Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that  
 binds and activates ErbbA." Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).  
 RL - FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR.  
 CC BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR,  
 CC ERBB2 OR ERBB3 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN SYMPATHETIC, MOTOR, AND SENSORY  
 CC NEURONS.  
 CC -1- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS E11. IN E13 EMBRYOS, DETECTED IN  
 CC THE BRAIN, SPINAL CORD, TRIGEMINAL, VESTIBULAR-COCHLEAR, AND  
 CC

Copyright (c) 1993 - 2000 Compugen Ltd.

1 nucleic - nucleic search, using sw model  
in on: January 8, 2002, 17:27:15 ; Search time 158.12 Seconds  
(without alignments)  
1886.853 Million cell updates/sec

title: US-09-640-041-3  
effect score: 348  
quence: 1 atggcaaacatgatccaga.....acatgtatggacacactga 348  
oring table: IDENTRY-NUC  
gapop 10.0 , Gapext 1.0  
arched: 930621 seqs, 424662619 residues  
otal number of hits satisfying chosen parameters: 1861242  
nimum DB seq length: 0  
ximum DB seq length: 200000000  
rst-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : N\_Geneseq\_1101:\*

1: /\$IDS2/gcgdata/geneseq/geneseq/NA1980.DAT:\*

2: /\$IDS2/gcgdata/geneseq/geneseq/NA1981.DAT:\*

3: /\$IDS2/gcgdata/geneseq/geneseq/NA1982.DAT:\*

4: /\$IDS2/gcgdata/geneseq/geneseq/NA1983.DAT:\*

5: /\$IDS2/gcgdata/geneseq/geneseq/NA1984.DAT:\*

6: /\$IDS2/gcgdata/geneseq/geneseq/NA1985.DAT:\*

7: /\$IDS2/gcgdata/geneseq/geneseq/NA1986.DAT:\*

8: /\$IDS2/gcgdata/geneseq/geneseq/NA1987.DAT:\*

9: /\$IDS2/gcgdata/geneseq/geneseq/NA1988.DAT:\*

10: /\$IDS2/gcgdata/geneseq/geneseq/NA1989.DAT:\*

11: /\$IDS2/gcgdata/geneseq/geneseq/NA1990.DAT:\*

12: /\$IDS2/gcgdata/geneseq/geneseq/NA1991.DAT:\*

13: /\$IDS2/gcgdata/geneseq/geneseq/NA1992.DAT:\*

14: /\$IDS2/gcgdata/geneseq/geneseq/NA1993.DAT:\*

15: /\$IDS2/gcgdata/geneseq/geneseq/NA1994.DAT:\*

16: /\$IDS2/gcgdata/geneseq/geneseq/NA1995.DAT:\*

17: /\$IDS2/gcgdata/geneseq/geneseq/NA1996.DAT:\*

18: /\$IDS2/gcgdata/geneseq/geneseq/NA1997.DAT:\*

19: /\$IDS2/gcgdata/geneseq/geneseq/NA1998.DAT:\*

20: /\$IDS2/gcgdata/geneseq/geneseq/NA1999.DAT:\*

21: /\$IDS2/gcgdata/geneseq/geneseq/NA2000.DAT:\*

22: /\$IDS2/gcgdata/geneseq/geneseq/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult	No.	Score	Query	Match	Length	DB	ID	Description
1	348	100.0	348	22	AAH91294			Human EGFR2 coding
2	252	72.4	348	22	AAH91295			Mouse EGFR2 coding
3	32	9.2	3663	18	AAH72320			Embryonic stem cell
4	31.6	9.1	383	22	AAH64385			Novel human polyu
5	30.6	8.8	1559	18	AAH91855			DUB-1 enhancer/pro
6	30.6	8.8	92407	22	AAH28549			Genomic fragment#
7	30.4	8.7	803	22	AAH22969			DNA encoding novel
8	30.4	8.7	4487	7	AAN0943			Sequence encoding
9	30.2	8.7	783	21	AAH33642			Arabidopsis thalia
10	29.8	8.6	1202	22	AAH46934			Human secreted pro
11	29.8	8.6	1947	22	AAH46950			Human secreted pro

## SUMMARIES

## ALIGNMENTS

OC	Mammalia; Buteraria; Cetartiodactyla; Suina; Suidae; Sus
NCBI_TAXID	9823;
RN	[1]
RP	SEQUENCE FROM N A.
RA	Kim J.G., Vallet J.L., Christenson R.K.;
RT	"Characterization of uterine epidermal growth factor during pregnancy in pigs."
RT	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF336151; AAKB830.1; -.
FT	CHAIN 970 1022 EPIDERMAL GROWTH FACTOR.
SQ	SEQUENCE 1214 AA; 133505 MW; 83D104057458CF CRC64;
Query	Match
Best	Local Similarity 20.6%; Score 128.5; DB 6; Length 1214;
Matches	Local Similarity 33.7%; Pred. No. 6; 2e-06; Indels 9; Gaps 2;
OY	9 CGPSHKSFCLNGGLCVIPTIPSPFRCVNTGARCEEVFLP-----GSS10TKNLF 63
Db	975 CPPSHDGICLHGGVCMYEAVDSIYACNCVFGVYGRGRCQHDLKWWELRHAGLGRQWNTV 1034
OY	64 AFVALAVLVLTVL---GAFYFLCRK 85
Db	1035 VAVCVVWVLVLULLGWGAHYRTQ 1060
RESULT	3
O9ESA9	PRELIMINARY; PRT; 298 AA..
ID	O9ESA9; PRELIMINARY; PRT; 298 AA..
AC	O9ESA9; PRELIMINARY; PRT; 298 AA..
DT	01-MAR-2001 (TREMBL). 16. Last sequence update)
DT	01-JUN-2001 (TREMBL). 17. Last annotation update)
DE	SMDF NEUREGULIN ALPHA 2B (FRAGMENT).
GN	NRG1.
OS	Rattus norvegicus (Rat);
OC	Batrachota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OX
NCBI_TaxID	10116;
RN	[1]
RP	SEQUENCE FROM N A.
RC	STRAIN=BDIX;
RA	Carroll S.L., Anderson K.D., Frohnhert P.W.;
RT	"Structural and Functional Diversity of SMDF Neuregulin Splice Variants Expressed in the Adult Rat Nervous System.;"
RR	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF194440; AAG28429.1; -.
DR	InterPro; IPR00561; EGF-like.
DR	InterPro; IPR002154; Neuregulin.
DR	InterPro; IPR002114; PTS_HPr_ser.
DR	Pfam; PF0008; EGF; 1.
DR	Pfam; PF02158; Neuregulin; 1.
DR	PRINTS; PRO1089; NEUREGULIN.
DR	SMART; SM0018; EGF; 1.
DR	SMART; SM00011; EGF-like; 1.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR	PROSITE; PS00589; PTS_HPr_SER; UNKNOWN_1.
FT	NON-TER 1
FT	NON-TER 298
SQ	SEQUENCE 298 AA; 32851 MW; BD76F014C2B33026 CRC64;
Query	Match
Best	Local Similarity 19.8%; Score 123.5; DB 11; Length 298;
Matches	Local Similarity 25.0%; Pred. No. 5; 3e-06; Indels 25; Gaps 6;
OY	3 TDHEECGSPSKSFCLNGGLCVIPTIPSPFRCVNTGARCEEVFLPSS10KS 59
Db	46 TSHLTKCAEKFTFCVNGGCFTVKOLNSRSYRLCKCQPG-TGARTE-NVP-MKVTQE-103
OY	60 NLFEATVALAVLVLTVL---LITGAFYFLCRK-----RKGHFQRASSVVOYDNLW 102
Db	104 KAELYQKRVITIGCIALLVG--IMCVVAYCKTKKORQKLHDLRSLRSERSNLV 160



Novel epidermal growth factor H2 (EGFH2) polypeptides and polynucleotides encoding them useful for treating EGFH2 protein-modulated disorders such as cancers, Alzheimer's disease, ischaemic stroke and epilepsy

DR WPI; 2001-211304/21.  
DR N-FSDB; AAA91295.

PT Novel epidermal growth factor H2 (EGF<sub>H2</sub>) polypeptides and  
PT polynucleotides encoding them useful for treating EGF<sub>H2</sub>  
PT protein-modulating disorders such as cancers, Alzheimer's disease,  
PT ischaemic stroke and epilepsy

XX

PS

XX

CC This sequence is the mouse epidermal growth factor H2 (EGF<sub>H2</sub>)  
CC protein of the invention. The EGF<sub>H2</sub> DNA and protein sequences are useful  
CC for modulating the amount of EGF<sub>H2</sub> protein in a subject, or treating an  
CC EGF<sub>H2</sub> protein modulated disorder. The DNA or its antisense construct,  
CC ribozyme or a retroviral vector comprising a promoter are also useful  
CC for treating breast, prostate, pancreatic, oral or ovarian cancers,  
CC peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's  
CC disease, Parkinson's disease, Huntington's disease, ischaemic stroke,  
CC brain injury, acute spinal cord injury, nervous system injury, multiple  
CC sclerosis, infection, dementia, epilepsy, peripheral nerve injury,  
CC acoustic trauma and tissue wound. The DNA and protein are also useful for  
CC providing trophic support for cells in a patient by implanting cells  
CC expressing them encapsulated in a semipermeable membrane. EGF<sub>H2</sub>  
CC polypeptides can be used to screen peptide libraries. Antibodies which  
CC bind to EGF<sub>H2</sub> and/or variant polypeptides can be used in diagnosing and  
CC determining the prognosis of cancer, tumour progression,  
CC hyperproliferative cell growth or accompanying biological and physical

Query	Best Match	Length	Score
QY	1 MPTHEEPPGPHSFLCNGLCVYIPTPSPFCRCVENTGARCEEVFPGSSTQTKSN	60	100 %; Score 625; DB 22; Pred. No. 1. 9e-51; Mismatches 0; Indels 0; Gaps 0;
Db	1 mptcheeppgphsfcflnglcgqyiptspfcrcvntgarceevfipgssstqtksn	60	
QY	61 LPEAFVALAVLVTLLIGAFYFLCRKGHQDASSVOYDINIVETSTSAAHHSHEQH	115	
Db	61 lfeafvalavltlligafyflcrkgfgrassvqydnivetsstsahnsheqh	115	
RESULT 2			
AYY97698	AYY97698 standard; Protein; 115 AA.		
XX			
AC	AYY97698;		
XX			
DT	19-JUN-2001 (first entry)		
XX			
DE	Mouse EGFR2 protein sequence.		
XX			
KW	EGFR2; epidermal growth factor H2; EGFR2 protein modulated disorder; cancer; peripheral neuropathy; amyotrophic lateral sclerosis; therapy; Alzheimer's disease; Parkinson's disease; Huntington's disease; ischaemic stroke; brain injury; acute spinal cord injury; infection; nervous system injury; multiple sclerosis; dementia; epilepsy; peripheral nerve injury; acoustic trauma; mouse; mus mus.		
XX			
PN	WO200114415-A2.		
XX			
PD	01-MAR-2001.		
XX			
PF	15-AUG-2000; 2000WO-US22326.		
PR	20-AUG-1999; 99US-0149986..		
XX			
PA	(CHIR ) CHIRON CORP.		
XX			
PI	Kavanaugh WM, Cen H, Lee P;		
XX			

Result No.	Score	Query Match	Length	DB	ID	Description
						GenCore version 4.5
						Copyright (c) 1993 - 2000 Compugen Ltd.
						nucleic - nucleic search, using sw model
on:		January 8, 2002, 15:14:05		Search time 1410.3 Seconds		
				(without alignments)		
				4070.778 Million cell updates/sec		
le: effect score: 348	1	348	100.0	348	6	AX088172
juence: 1 atgccaacagatcacaaga.....acagtcatgaacacactga	2	252	72.4	348	6	AX088170
ring table: IDENTITY_NUC	3	252	72.4	750	10	AF083067
gap0 10.0 , Gapext 1.0	4	157.8	45.3	166352	2	AC026180
ximum DB seq length: 0	5	157.8	45.3	171000	2	AC087456
ximum DB seq length: 200000000	6	157.8	45.3	189076	2	AC068838
st-processing: Minimum Match 0%	7	101.2	29.1	154323	2	AC026033
Maximum Match 100%	8	101.2	29.1	161262	2	AC087464
Listing first 45 summaries	9	101.2	29.1	166352	2	AC026180
tabase : Genbank:*	10	90	25.9	180795	2	AC027104
1: gb_ba:*	11	48.6	14.0	188554	2	AC019618
2: gb_htg:*	12	37.2	10.7	204779	2	CNS01RH3
3: gb_in:*	13	35.4	10.2	2691	4	BOVBUTBT1
4: gb_om:*	14	35.4	10.2	170139	2	AC073727
5: gb_ov:*	15	35	10.1	190577	2	AC073721
6: gb_pat:*	16	35	10.1	199694	9	AC005036
7: gb_ph:*	17	35	10.1	199694	9	AC005036
8: gb_pl:*	18	35	10.1	24174	4	AC09564
9: gb_pr:*	19	34.8	10.0	46326	5	AF10420
10: gb_ro:*	20	34.6	9.9	155084	9	AL136220
11: gb_sts:*	21	34.6	9.9	154242	2	AC073737
12: gb_sy:*	22	34.2	9.8	1121	14	AB023558
13: gb_un:*	23	34.2	9.8	128294	9	AC07248
14: gb_vl:*	24	34.2	9.8	163117	9	AC018733
15: em_ba:*	25	33.8	9.7	2855	8	SPU73044
16: em_fun:*	26	33.8	9.7	23851	8	SPB1861
17: em_hum:*	27	33.6	9.7	151798	2	AC044819
18: em_in:*	28	33.6	9.7	159991	2	AC069199
19: em_om:*	29	33.6	9.7	236810	2	AL162734
20: em_or:*	30	33.4	9.6	37968	8	CGC20C1
21: em_ov:*	31	33.4	9.6	75382	2	AC023242
22: em_pat:*	32	33.2	9.5	224863	2	AC093165
23: em_ph:*	33	33	9.5	177767	2	AC097342
24: em_pl:*	34	33	9.5	179202	2	AC025935
25: em_ro:*	35	33	9.5	188863	9	AC007345
26: em_sts:*	36	33	9.5	190963	2	AC023597
27: em_sy:*	37	32.8	9.4	18667	2	AC092170
28: em_vl:*	38	32.8	9.4	36654	2	AC017551
29: em_vl:*	39	32.8	9.4	126908	2	AC012365
30: em_htg_hum:*	40	32.8	9.4	139515	2	AP002432
31: em_htg_inv:*	41	32.8	9.4	159804	2	AC055876
32: em_htg_rnd:*	42	32.8	9.4	169152	9	AP000793
33: em_htg_hum:*	43	32.8	9.4	17740	3	AC005714
34: em_htg_inv:*	44	32.8	9.4	181893	2	AC028901
35: em_htg_rnd:*	45	32.8	9.4	191466	3	AC008348
36: em_htg_other:*						
						AL160314 Homo sapi
						AC073551 Bovine buty
						AC073727 Mus muscu
						AL136220 Human DNA
						AL155206 Human ade
						AB023558 Human ade
						AC073721 Mus muscu
						AC005036 Homo sapi
						AC019614 Homo sapi
						AC071973 Schizosacch
						U73044 Schizosacch
						AL109814 S.pombe
						AC004819 Homo sapi
						AC073721 Homo sapi
						AC069199 Homo sapi
						AC073721 Homo sapi
						AC0162734 Homo sapi
						AI033391 C.albican
						AC023242 Homo sapi
						AC093165 Homo sapi
						AC007342 Homo sapi
						AC025935 Homo sapi
						AC007345 Homo sapi
						AC023597 Homo sapi
						AC092170 Homo sapi
						AC017551 Drosophila
						AC012365 Homo sapi
						AP002432 Homo sapi
						AC023597 Homo sapi
						AC000793 Drosophila
						AC005714 Drosophila
						AC019014 Homo sapi
						AC008348 Drosophili
						AL160314 Homo sapi
						AC073551 Bovine buty
						AC073727 Mus muscu
						AL136220 Human DNA
						AL155206 Human ade
						AB023558 Human ade
						AC073721 Mus muscu
						AC005036 Homo sapi
						AC019614 Homo sapi
						AC071973 Schizosacch
						U73044 Schizosacch
						AL109814 S.pombe
						AC004819 Homo sapi
						AC073721 Homo sapi
						AC069199 Homo sapi
						AC073721 Homo sapi
						AC0162734 Homo sapi
						AI033391 C.albican
						AC023242 Homo sapi
						AC093165 Homo sapi
						AC007342 Homo sapi
						AC025935 Homo sapi
						AC007345 Homo sapi
						AC023597 Homo sapi
						AC092170 Homo sapi
						AC017551 Drosophili
						AC012365 Homo sapi
						AP002432 Homo sapi
						AC023597 Homo sapi
						AC000793 Drosophili
						AC005714 Drosophili
						AC019014 Homo sapi
						AC008348 Drosophili
						AL160314 Homo sapi
						AC073551 Bovine buty
						AC073727 Mus muscu
						AL136220 Human DNA
						AL155206 Human ade
						AB023558 Human ade
						AC073721 Mus muscu
						AC005036 Homo sapi
						AC019614 Homo sapi
						AC071973 Schizosacch
						U73044 Schizosacch
						AL109814 S.pombe
						AC004819 Homo sapi
						AC073721 Homo sapi
						AC069199 Homo sapi
						AC073721 Homo sapi
						AC0162734 Homo sapi
						AI033391 C.albican
						AC023242 Homo sapi
						AC093165 Homo sapi
						AC007342 Homo sapi
						AC025935 Homo sapi
						AC007345 Homo sapi
						AC023597 Homo sapi
						AC092170 Homo sapi
						AC017551 Drosophili
						AC012365 Homo sapi
						AP002432 Homo sapi
						AC023597 Homo sapi
						AC000793 Drosophili
						AC005714 Drosophili
						AC019014 Homo sapi
						AC008348 Drosophili
						AL160314 Homo sapi
						AC073551 Bovine buty
						AC073727 Mus muscu
						AL136220 Human DNA
						AL155206 Human ade
						AB023558 Human ade
						AC073721 Mus muscu
						AC005036 Homo sapi
						AC019614 Homo sapi
						AC071973 Schizosacch
						U73044 Schizosacch
						AL109814 S.pombe
						AC004819 Homo sapi
						AC073721 Homo sapi
						AC069199 Homo sapi
						AC073721 Homo sapi
						AC0162734 Homo sapi
						AI033391 C.albican
						AC023242 Homo sapi
						AC093165 Homo sapi
						AC007342 Homo sapi
						AC025935 Homo sapi
						AC007345 Homo sapi
						AC023597 Homo sapi
						AC092170 Homo sapi
						AC017551 Drosophili
						AC012365 Homo sapi
						AP002432 Homo sapi
						AC023597 Homo sapi
						AC000793 Drosophili
						AC005714 Drosophili
						AC019014 Homo sapi
						AC008348 Drosophili
						AL160314 Homo sapi
						AC073551 Bovine buty
						AC073727 Mus muscu
						AL136220 Human DNA
						AL155206 Human ade
						AB023558 Human ade
						AC073721 Mus muscu
						AC005036 Homo sapi
						AC019614 Homo sapi
						AC071973 Schizosacch
						U73044 Schizosacch
						AL109814 S.pombe
						AC004819 Homo sapi
						AC073721 Homo sapi
						AC069199 Homo sapi
						AC073721 Homo sapi
						AC0162734 Homo sapi
						AI033391 C.albican
						AC023242 Homo sapi
						AC093165 Homo sapi
						AC007342 Homo sapi
						AC025935 Homo sapi
						AC007345 Homo sapi
						AC023597 Homo sapi
						AC092170 Homo sapi
						AC017551 Drosophili
						AC012365 Homo sapi
						AP002432 Homo sapi
						AC023597 Homo sapi
						AC000793 Drosophili
						AC005714 Drosophili
						AC019014 Homo sapi
						AC008348 Drosophili
						AL160314 Homo sapi
						AC073551 Bovine buty
						AC073727 Mus muscu
						AL136220 Human DNA
						AL155206 Human ade
						AB023558 Human ade
						AC073721 Mus muscu
						AC005036 Homo sapi
						AC019614 Homo sapi
						AC071973 Schizosacch
						U73044 Schizosacch
						AL109814 S.pombe
						AC004819 Homo sapi
						AC073721 Homo sapi
						AC069199 Homo sapi
						AC073721 Homo sapi
						AC0162734 Homo sapi
						AI033391 C.albican
						AC023242 Homo sapi
						AC093165 Homo sapi
						AC007342 Homo sapi
						AC025935 Homo sapi
						AC007345 Homo sapi
						AC023597 Homo sapi
						AC092170 Homo sapi
						AC017551 Drosophili
						AC012365 Homo sapi
						AP002432 Homo sapi
						AC023597 Homo sapi
						AC000793 Drosophili
						AC005714 Drosophili
						AC019014 Homo sapi
						AC008348 Drosophili
				</td		

RESULT 3  
US-08-168-091A-2  
; Sequence 2, Application US/08168091A  
; Patent No. 5665862  
; GENERAL INFORMATION:  
; APPLICANT: Flischbach, Gerald.  
; APPLICANT: Falls, Douglas R.  
; APPLICANT: Rosen, Kenneth M.  
; APPLICANT: Corras, Gabriel.  
; TITLE OF INVENTION: Neurotrophic Factor  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE AND COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/168,091A  
; FILING DATE: 15-DEC-1993  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/953,744  
; FILING DATE: 29-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deconti, Giulio A  
; REFERENCE/DOCKET NUMBER: HMI-002CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 602 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-168-091A-2  
; US-09-126-121-2  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 713 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; FEATURE:  
; NAME/KEY: Mouse NRG3 (mNRG3)/amino acid seq.  
; LOCATION: 1-713  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION:  
; US-09-126-121-2

Query Match 22.7%; Score 142; DB 1; Length 602;  
Best Local Similarity 22.1%; Pred. No. 8.3e-08;  
Matches 35; Conservative 21; Mismatches 51; Indels 22; Gaps 4;

Qy 3 TDHBEP CGPSHKSPCLNGGICVYIPTIPS--PFCRCVENYTGARCEEV----- 48  
Db 135 TSHLTKCDIKARCKCNGECIMWKDLPNPVRYLCRCPNEFTGRCQVMASEYKHLI 194  
Qy 49 -FLPGS10TSKSNLFEAFALAVLVTLLIGAFYFLC-----RIGHFORASSYQYDILV 101  
Db 195 EFMERAEIYQKRVLTITGICIAILVVGIMCVVAY-CKTKKQRKKLHDRLQSLSERNNV 253

RESULT 4  
US-08-428-926-5  
; Sequence 5, Application US/08428926  
; Patent No. 5667780  
; GENERAL INFORMATION:  
; APPLICANT: Ho, Wei-Hsien  
; APPLICANT: Osheroff, Phyllis L.  
; TITLE OF INVENTION: SENSORY AND MOTOR NEURON DERIVED FACTOR (SMDF)  
; NUMBER OF SEQUENCES: 5

Qy 52 ---GSS10TSKSNLFEAFALAVLVTLLIGAFYFLC-----RIGHFORASSYQYDILV 95  
Db 345 TDHLGIEFMEEDVYQKVISICIFGIVIVGMCAAFYFKSKK---QAKQIQEHLKE 400  
Qy 96 -----YDINLVEETSTSAHSH 112  
Db 401 SONGKNYSLKASSTKSESLMKSH 423